HELIX-Atlanta Benchmark Report:

The Application of Earth Science Satellite Observations to Improve Environmental Public Health Surveillance Systems

Partners: NASA/Marshall Space Flight Center, Centers for Disease Control and Prevention, U.S. Environmental Protection Agency, Kaiser Foundation Health Plan of Georgia, Inc, Georgia Environmental Protection Division, Georgia Institute of Technology, and Emory University

September 2005

Prepared by Marshall Space Flight Center and the Universities Space Research Association in Cooperation with the Centers for Disease Control and Prevention and Partners

Authors: NASA, Dr. Douglas Rickman, Dr. Dale Quattrochi, and Dr. Jeffrey Luvall

USRA, Dr. William Crosson, Mr. Maurice G. Estes, Jr., Dr. Ashutosh Limaye, and Dr. Mohammad Al-Hamdan

TABLE OF CONTENTS

Tables

Table 1. Linear correlation coefficients by month and sensor, and regression coefficients for April-September for each year and sensor. $PM_{2.5}$ is the dependent variable and AOD is the dependent variable: $PM_{2.5}$ = Slope*AOD + Intercept.

Figures

1.0 Executive Summary

The Health and Environment Linked for Information Exchange (HELIX)-Atlanta project is an effort between the NASA Marshall Space Flight Center (NASA/MSFC) and the National Center for Environmental Health (NCEH) at the Centers for Disease Control and Prevention (CDC) to integrate NASA earth science satellite observations and modeling capabilities into an Environmental Public Health Tracking Network (EPHTN) for five counties in the Atlanta metropolitan area. The EPHTN exists as an active and evolving Decision Support System (DSS) to enable researchers and public health officials to collect, integrate, analyze, interpret, maintain, and disseminate data on environmental hazards, human exposure to those hazards, and potentially environmentally-related diseases. Congress initiated funding for the Environmental Public Health Tracking (EPHT) program because it recognized that the nation lacked critical data that could document possible links between environmental hazards and noninfectious diseases or other health effects. The focus of EPHT is to improve the health of communities. Information from an EPHTN will help federal, state and local agencies be better prepared to develop and evaluate effective public health actions to prevent or control diseases or other disabilities that are linked to hazards in the environment. Additionally, health care providers will be able to provide better care and targeted preventive services, and the public will have a better understanding of what is happening in their communities and what actions they can take to protect or improve their health. As a key effort within the NASA Applied Sciences Program's Public Health National Application, the Health and Environment Linked for Information Exchange in Atlanta project (HELIX-Atlanta) is focusing on activities that can provide information for:

- Evaluating the effectiveness of environmental public health interventions
- Facilitating environmental public health decision making
- Monitoring and detecting changes of selected environmental public health events
- Developing environmental public health hypotheses for research

Within this purview, NASA/MSFC and the CDC along with other partners are working to examine how remote sensing and other geospatial data can provide definitive information that can be of utility to the overall functionality of an EPHTN. Since the inception of HELIX-Atlanta in 2003, NASA/MSFC has worked with the CDC to define the data parameters needed for an EPHTN over the Atlanta metropolitan study area, helped to establish the overall design for incorporating geospatial data into the EPHTN, and has provided technical expertise in using NASA earth science satellite observations and modeling capabilities to estimate aerosol particulate matter (i.e., Particulate Matter [PM] of 2.5 microns in size or smaller [$PM_{2.5}$]). NASA/MSFC has the ability to estimate $PM_{2.5}$ for the greater Atlanta area and then match these data to data on clinical visits for asthma exacerbations, which has been vital to the overall success of HELIX-Atlanta. NASA/MSFC's ability to integrate MODIS aerosol optical depth observations with data available from the U.S. Environmental Protection Agency's ground-level PM monitoring stations is of significant importance to the CDC because it is in need of such data for public health surveillance within a EPHTN. This report describes the rationale for the HELIX-Atlanta project, the overall design of the DSS as related to CDC's needs, and the flow of information from public health and environmental data sources to the HELIX-Atlanta EPHTN.

2.0 Introduction

2.1 NASA Application Traceability and Link to NASA Mission

As part one of the twelve National Priority themes of the NASA Applied Sciences Program, the Public Health application extends products derived from earth science information, models, technology, and other capabilities into decision support tools for public health, medical and environmental health issues. This application draws on Earth observations such as weather, climate, natural hazards, and other environmental factors to address public health, medical, and environmental health issues (NASA, 2005). The Public Health application focuses on partnerships with the public health practice community and decision support systems related to Epidemiologic Surveillance Systems in the areas of infectious disease, environmental health, and bioterrorism.

NASA collaborates with the professional public health community that is responsible for surveillance to understand and respond to factors in the environment that adversely impact the health of the American public. Within this overall purview, the Public Health application theme embodies the principles related to a Decision Support Systems (DSS) concept that is the fulcrum for the NASA Applied Sciences program. To initiate a Public Health DSS, an Integrated System

Solution (ISS) construct exists as the foundation for the DSS. One of the principal segments of the Public Health DSS is Environmental Public Health Tracking (EPHT), the ongoing collection, integration, analysis, and interpretation of data about factors related to environmental hazards, exposure to environmental hazards, and health effects potentially related to exposure to environmental hazards. The goal of EPHT is to protect communities by providing information to federal, state, and local agencies. These agencies in turn, use this information to plan, apply, and evaluate public health actions to prevent and control environmentally related diseases.

2.1.1 Link to Applied Sciences Roadmap and IWGEO

The roadmap for the Public Health National Application theme defines the direction, identifies key factors, and identifies and communicates the evolutionary path to reach this theme's objectives. The HELIX-Atlanta project is a critical element in the roadmap as associated with bringing NASA earth science data standards, workforce development support, and sustained integration of satellite observations and models into the DSS construct to develop a public health surveillance system that potentially will automatically ingest data and models to augment the EPHT DSS that will be used by the CDC. Thus, HELIX-Atlanta constitutes a vital component of the NASA Public Health National Application and its goal of transcending predictions and NASA earth science satellite observations into DSS tools and translating the DSS outputs into values and benefits to reduce environmentally-related diseases for the U.S. public.

Additionally, HELIX-Atlanta directly relates to the overall framework of the Interagency Working Group on Earth Observations (IWGEO) and the Global Earth Observation System of Systems (GEOSS) by being associated with one of the nine societal benefit areas established by IWGEO, specifically in "Understanding environmental factors affecting human health and well-being" (GEOSS, 2005). Moreover, EPHT and HELIX-Atlanta are integral to the success of IWGEO in attaining the goals established for the program's 2 year, 6 year, and 10 year targets to:

- Promote the development of an integrated public health information network database that includes information relevant to human health officials and agencies, and includes multi-scaled, multi-temporal spatial data collected from satellite observations sources, to provide better predictive models of the effects of environmental factors affecting human health and well-being (2 year target).

- Facilitate development of data products and systems that integrate Earth science databases with health and epidemiological information. This includes social and infrastructure data needed in decision support systems for health care planning and delivery (2 year target).
- Produce an inventory of available Earth remote sensing and ground-based databases that can be associated with known health problems such as asthma, pollutant exposure, birth defects, seafood contamination and certain infectious and vector-borne diseases. This includes remote sensing and ground-based databases, historic data sets encompassing well characterized epidemics, and gaps in human health related environmental data (e.g., places where water, soil, or air quality are not measured). To accomplish this, GEOSS will develop the tools, architecture and infrastructure for a public health community at large to obtain historical and current health data for better predictability of environmental effects on human health (6 year target).
- Facilitate the early detection and control of environmental risks to human health through improvements in the sharing and integration of Earth observations, monitoring, and early warning systems, databases, models, and communications systems (10 year target).
- Advocate better on-ground disease surveillance, linked with open national reporting practices, for better understanding and documentation of environmental influences on infections, chronic and other diseases and disorders (10 year target) (GEOSS, 2005).

2.2 Brief Introduction to Decision Support System

HELIX-Atlanta is an active and evolving DSS. Statutory responsibility for developing and managing HELIX-Atlanta, a prototype for the national EPHTN, belongs to the CDC. A bill to create the national EPHTN entitled, "Nationwide Health Tracking Act of 2002" was introduced in the 107th Congress. The EPHTN will establish a national network of local, state, and federal public health agencies to track trends in priority non-infectious health effects. Around 2009, when fully functional, the national EPHTN will be an early warning system for the rapid identification of adverse health events such as toxic chemical releases. The national EPHTN also will integrate data on exposure to environmental hazards to assist in characterizing correlations between adverse health effects and the environment. Earth science results will provide new information on the environmental contribution to non-infectious diseases and other adverse health effects and predictive value based on coupled Earth system-non-infectious health effect models.

The Environmental Health Tracking Branch, Division of Environmental Hazards and Health Effects, National Center for Environmental Health (NCEH) within CDC is developing HELIX-Atlanta, to demonstrate a process for developing a local environmental public health tracking (surveillance) network that integrates non-infectious health and environment information systems. HELIX-Atlanta will inform national and international efforts. For more background see McGeehin, et al., 2004.

3.0 Systems Engineering Approach

Figure 1 illustrates the architecture that is the fulcrum for the Applied Sciences Program as related to an Integrated System Solutions construct. The flow of the process as indicated in the figure is such that NASA as a research and development agency extends the observations, model predictions, and computational techniques from earth science research to support partner agencies. The right side of the diagram relates to the engagement of partner agencies and organizations that develop and operate Decision Support Tools (DSTs) to analyze scenarios, identify alternatives, and assess risks as part of their respective decision-making processes. Federal agencies, for example, use these tools to support their responsibilities to the public, such as resource management, security, regulations, public health, and economic development. In the middle of the diagram are DSTs. DST here refers to assessments and decision support systems that serve policy and management decisions. Typically, DSSs are interactive, computer-involved systems that are used to retrieve information, analyze alternatives, and evaluate scenarios to gain insight into critical factors, sensitivities, and consequences of potential decisions. DSSs as inputs and synthesizers of great quantities of earth science data and computationally demanding scientific models exist as systematic mechanisms to incorporate data products and document the value derived from inputs into the DSS process. In this project, the evolving Environmental Public Health Surveillance DSS is enhanced with NASA earth science satellite observations to improve the quantity and quality of environmental measurements and spatial coverage of the five-county HELIX-Atlanta pilot study area. These data are expected to improve the overall function of the DSS and support the ultimate goal of enabling better public health policy and management decisions.

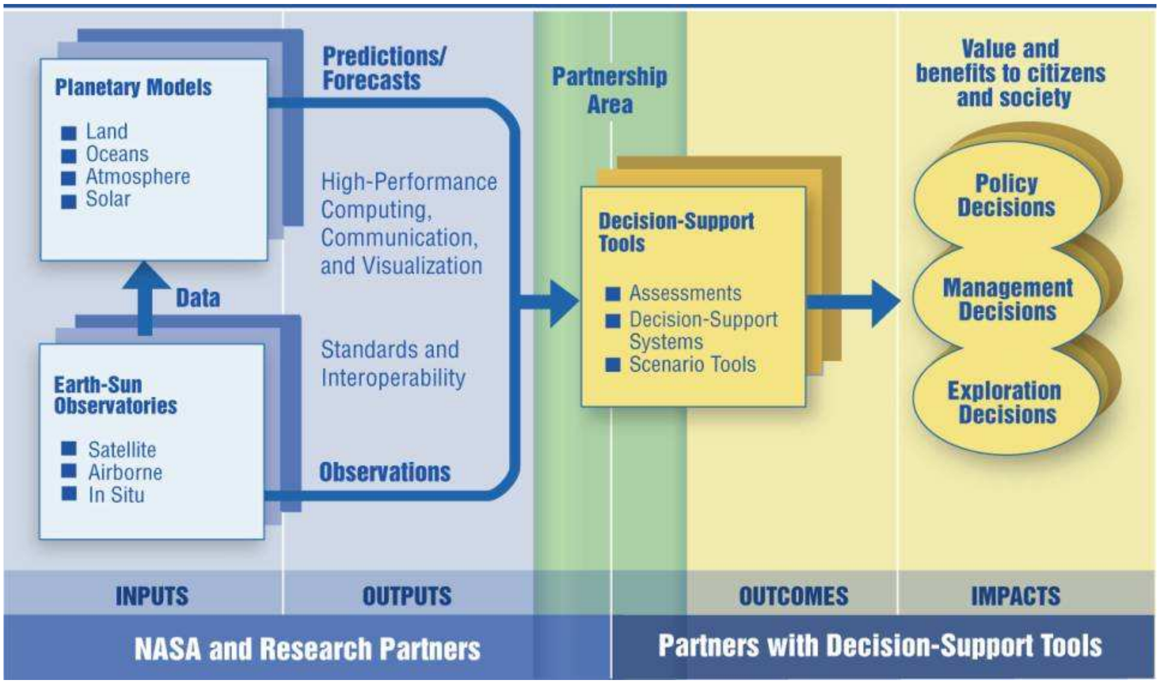


Figure 1. Flow diagram of the Integrated System Solutions Architecture

The Integrated System Solution Architecture process employs a systems engineering approach to achieve both functionality and consistency across a variety of earth science applications. Embedded within this systems engineering schema are critical junctions in the process relating to Evaluation, Verification and Validation (V&V), and Benchmarking (http://aiwg.gsfc.nasa.gov/guidelines.html). The Evaluation phase provides an initial assessment of user-defined requirements relative to earth science research results. Typically this includes identifying decision-support tools associated with an application area, assessing the potential value and technical feasibility of current and future earth science results within the purview of a DST concept, and assessing partner commitment and project value versus benefit relative to the Applied Sciences Program's funding and objectives. As part of Evaluation from a DSS perspective, a decision is made as part of this phase to continue on with collaboration into the Verification and Validation phase.

V&V focuses on developing prototype products to address requirements, devise system integration approaches, and resolve technical issues related to the introduction of the earth science products into DSTs. Paramount in this systems engineering phase is the measurement and performance characteristics of earth science products (i.e., NASA outputs) to meet the input requirements of the DSTs by addressing issues associated with bringing satellite observations and model outputs into the partners' internal systems. Here the ultimate purpose of V&V is to ensure that the end-to-end system meets the intended objectives with the new inputs; i.e., that the system functions properly given the ingestion of new earth science satellite observations and models. Verification determines how the actual performance of an observation, prediction, or other earth science product meets the user-defined requirements within a specific tolerance range. Validation determines if the performance of the algorithms (or logic or rationale) using earth science satellite observations or models can accomplish the stated intended outcomes.

The Benchmarking phase involves the application of a rigorous process to compare the performance of a DST using earth science products to a standard, recognized criterion or set of criteria, as well as a current practice or reference scenario to document the value of earth science products within the scope of the DST. If partners have existing metrics and performance standards to evaluate their tools and decisions, these are used as the standard within the Benchmarking phase. The robust documentation of procedures and guidelines that describe the steps to access and utilize earth science satellite observations and results is a requisite part of this systems engineering phase.

The three phases embodied within the systems engineering concept provide a systematic approach to follow the integrated systems solutions architecture and apply NASA's satellite observations and models within an Integrated Systems Solutions methodology. This report documents our efforts at improving a Public Health Surveillance System using the process flows and robustness defined by the NASA Applied Sciences Program's Integrated Systems Solutions Architecture.

3.1. Environmental Public Health Surveillance as a contrast to the NASA Systems Engineering model

HELIX-Atlanta follows standardized CDC practice in design and implementation of a public health surveillance program:

1. Establish objectives
2. Develop case (event) definitions
3. Determine data sources, data collection mechanisms and type of system
4. Develop data collection instruments
5. Field-test methods

6. Develop and test an analytic approach
7. Develop a dissemination mechanism
8. Assure use of analysis and interpretation

This modality is the analog of NASA systems engineering approach, but designed to meet the realities of epidemiology. HELIX-Atlanta teams are working on steps 1 – 5 at this time, varying from team to team. The project has five discrete teams working with existing information systems to answer public health practice inquiries from the public, policy-makers, and other stakeholders. The teams are Developmental Disabilities, Cancer, Birth Defects, Respiratory Health and Water.

3.2 Definition of Problem

Particulate Matter, ozone and other air pollutants pose a major health problem in the U.S. and are therefore regulated by the U.S. Environmental Protection Agency (EPA). It is the role of federal, state, and local agencies to develop plans and policies to comply with the National Ambient Air Quality Standards (NAAQS) and promote public health. Public health surveillance is defined as the ongoing, systematic collection, analysis, and interpretation of data on specific health events for use in the planning, implementation and evaluation of public health programs. Environmental hazards including air pollutants, water pollutants, and chemical agents such as lead and benzene among others are of interest in this project.

The purposes of the National Environmental Policy Act (NEPA) are: To declare a national policy which will encourage productive and enjoyable harmony between man and his environment; to promote efforts which will prevent or eliminate damage to the environment and biosphere and stimulate the health and welfare of man; to enrich the understanding of the ecological systems and natural resources important to the Nation; and to establish a Council on Environmental Quality.

The Clean Air Act requires the U.S. EPA to set NAAQS for common air pollutants. The standards are to be set at levels that protect public health with an adequate margin of safety to protect the health of sensitive groups of people. The standards drive the nation's air pollution control programs and must be reviewed every five years to ensure that public health issues are being adequately addressed. The Clean Air Act requires the states and EPA to develop strategies for reducing pollution from cars, factories and power plants in order to meet the air quality standards.

3.2.1 Particulate Matter

Particle pollution is a mixture of microscopic solids and liquid droplets suspended in air. This pollution, also known as particulate matter, is made up of a number of components, including acids (such as nitrates and sulfates), organic chemicals, metals, soil or dust particles, and allergens (such as fragments of pollen or mold spores). Fine particle pollution or $PM_{2.5}$ describes particulate matter that is less than or equal to 2.5 μm in diameter - 1/30th the diameter of a human hair. Smog, including particulate matter, is found in both major cities especially during summer months and in natural environments such as the Great Smoky Mountains (EPA: http://www.epa.gov/pmdesignations/basicinfo.htm).

Fine particle pollution can be emitted directly or formed secondarily in the atmosphere. Sulfates are a type of secondary particle formed from sulfur dioxide emissions from power plants and industrial facilities. Nitrates, are formed from emissions of nitrogen oxides from power plants, automobiles, and other combustion sources. The chemical composition of particles depends on location, time of year, and weather.

Health studies have shown a significant association between exposure to fine particles and premature death from heart or lung disease. Older adults and children with preexisting conditions are particularly vulnerable. Fine particles can aggravate heart and lung diseases and have been linked to effects such as cardiovascular symptoms, cardiac arrhythmias, heart attacks, respiratory symptoms, asthma attacks and bronchitis. These effects can result in increased hospital admissions, emergency room visits, absences from school or work, and restricted activity days.

Recently, the EPA began providing forecasts for $PM_{2.5}$. Particulates in this size range are called respirable aerosols and are easily entrapped by the lungs. Pollutants and diseases carried by respirable aerosols are a significant health threat (NASA Langley Research Center, Benchmark Report, 2003). According to the World Research Institute, an environmental research and policy organization, "the health effects of particulates are strongly linked to particle size. Small particles, such as those from fossil fuel combustion, are likely to be the most dangerous, because they can be inhaled deeply into the lungs, settling in areas where the body's natural clearance mechanisms can't remove them" (WRI, 1999).

EPA issued the fine particle standards in 1997 after evaluating hundreds of health studies and conducting an extensive peer review process. The annual standard is a level of 15 micrograms per cubic meter, based on the 3-year average of annual mean $PM_{2.5}$ concentrations. The 24-hour standard is a level of 65 micrograms per cubic meter, determined by the 3-year average of the annual 98th percentile concentrations (EPA). EPA is now engaged in the next five-year review of the health standards for particulate matter and ozone. (http://www.cleanairstandards.org/)

3.3 Description of DSS

3.3.1 Relevance and NASA interest in Public Health Applications

Under the auspices of the Public Health National Application NASA/MSFC is working with the CDC within the HELIX-Atlanta project to develop an EPHTN for five counties (Clayton, Cobb, DeKalb, Fulton, and Gwinnett) in the Atlanta, Georgia metropolitan area. This project seeks to develop an EPHTN to:

- Evaluate effectiveness of environmental public health interventions
- Facilitate environmental public health decision making
- Monitor and detect changes of selected environmental public health events
- Develop environmental public health hypotheses for research.

The HELIX-Atlanta network will be part of the national EPHTN. Initially, however, HELIX-Atlanta activities will focus on methods development for preparing health- and environment-related data for integration into a local EPHTN. These activities will result in increased EPHT capacity in the five-county area and provide credible information to advance public health practice and research. CDC's goal is to develop a tracking system that integrates data about environmental hazards and exposures with data about diseases that are possibly linked to the environment. This system as a DSS will allow federal, state, and local agencies and others to:

- Monitor and distribute information about environmental hazards and disease trends,
- Advance research on possible linkages between environmental hazards and disease,
- Develop, implement, and evaluate regulatory and public health actions to prevent or control environmentally related diseases.

From this perspective, HELIX-Atlanta as part of the Public Health National Application is integral to the DSS ISS in developing and demonstrating:

- Technologies for verification and validation of satellite observations to serve the EPHT network concept,
- Data pipelines for direct feed of environmental measurement data to EPHT-networked systems,
- New or improved environmentally-related disease models,
- Image output maps derived from NASA satellite and other spatial sources for EPHTN monitoring and scientific analysis.

3.3.2 CDC/NASA Partnership to Enhance HELIX-Atlanta

The CDC's collaboration with NASA/MSFC to develop methods useful for building the EPHTN is critical to the success of CDC's national EPHT Program. NASA scientists play an integral role in the EPHT HELIX-Atlanta project. Their work to develop tools to better characterize population exposure to $PM_{2.5}$ and to link exposure and health data in HELIX-Atlanta advances national and international efforts to utilize global earth observation systems for promotion of public health.

Meetings between NASA and CDC staff have provided opportunity for crucial knowledge transfer between the agencies. MSFC input has been essential for CDC to better understand the potential uses of satellite observations for EPHT efforts in the greater Atlanta area and potentially nationally. MSFC's knowledge of geostatistics has also been helpful in the CDC's efforts to translate observations from satellites into information that is meaningful for public health applications.

NASA provides the technical expertise in using observations from satellites to estimate $PM_{2.5}$ levels at the Earth's surface. Temporal and spatial gaps exist for *in situ* air monitoring data. NASA is working with CDC and other partners to examine how earth science satellite observations may fill these gaps and enhance our understanding of exposure of populations to $PM_{2.5}$. NASA's development and use of algorithms that will integrate Moderate Resolution Imaging Spectroradiometer (MODIS) aerosol optical depth observations and EPA air monitoring data to estimate $PM_{2.5}$ exposure is beyond CDC's technical ability. NASA/MSFC's ability to estimate $PM_{2.5}$ for the greater Atlanta area and then match that data to data on clinical visits for asthma exacerbations has been vital to the HELIX-Atlanta project. The Integrating of HMO

Respiratory Health Information with Air Pollution Data is ongoing. $PM_{2.5}$ exposure estimates generated by NASA scientists will also be used in a second HELIX-Atlanta project on Air Pollution and Congenital Health Defect in Atlanta.

This approach has significant potential for helping CDC track and assess particulate matter's relationship to human health in the greater Atlanta area and elsewhere in the country. Continued collaborations with NASA/MSFC is essential to advance CDC's pursuit of incorporating satellite observations with air quality data collected at EPA's air quality sites. This may offer improved coverage both temporally and spatially for particulate matter in ambient air and has the potential to add considerable power to studies that compare health effects with ambient air concentrations of $PM_{2.5}$. Further, it offers the prospect of studying the relationship between particulate matter in ambient air and health outcomes in regions which lack surface monitoring of particulate matter. Satellite observations may also prove to be very useful in tracking and assessing the risk of large $PM_{2.5}$ transport events such as forest fires to the public's health.

3.3.3 Flow of Information from Public Health and Environmental Data Sources to HELIX-Atlanta EPHT System

Two sources of $PM_{2.5}$ data were identified for use in HELIX-Atlanta. The first is the EPA Air Quality System (AQS) which measures many air pollutants at a network of ground monitoring stations across the U.S. (http://www.epa.gov/air/data/aqsdb.html). The AQS network provides $PM_{2.5}$ measurements from different monitoring stations concentrated around the metropolitan area and a few monitors in rural areas. AQS observations are made at time intervals ranging from one hour to six days. The second $PM_{2.5}$ data source is NASA's MODIS satellite which provides a measure of Aerosol Optical Depth (AOD), a measure of the degree to which sunlight is scattered and absorbed by aerosols of various sizes throughout the entire atmospheric column. AOD has been found to relate to ground level $PM_{2.5}$ under certain conditions. The MODIS AOD observations are at an approximate 10 km spatial resolution. MODIS data sets are available for each day of the year. Cloud cover prevents calculation of MODIS AOD and thus the calculation of $PM_{2.5}$.

Kaiser Foundation Health Plan of Georgia, Inc. (KP-GA) provides Protected Health Information (PHI) (i.e. acute asthma data) for the 5-county HELIX-Atlanta area to NASA scientists for linkage purposes under a Business Associate Agreement (BAA) between the two parties. KP-GA will send the PHI to NASA after applying an encryption procedure. After the health and environmental data are linked, NASA scientists will apply an additional encryption procedure and send the linked data set, which includes individual health information, back to KP-GA for further analysis. Also, the linked data will be aggregated into surfaces of 10 km by 10 km resolution to protect the privacy and confidentiality of KP-GA health members. The aggregated data set will be sent back to the CDC and published for the public.

3.3.4 PHIN, HIPPA, and Other Data Issues

The Public Health Information Network (PHIN) is CDC's vision for advancing fully capable and interoperable information systems in the many organizations that participate in public health. PHIN is a national initiative to implement a multi-organizational business and technical architecture for public health information systems. With the acceptance of information technology as a core element of public health, public health professionals are actively seeking essential tools capable of addressing and meeting the needs of the community. PHIN includes a portfolio of software solutions and artifacts necessary in building and maintaining interconnected information systems throughout public health at the local, state and federal levels (CDC). The MSFC team has participated in the Public Health Architecture working group since its inception and has provided input to partner members regarding the specialized requirements of satellite image accession and transference over public health networks. The team has also provided extensive discussion and expertise regarding Geographic Information Systems (GIS) and large dataset utilization. These are significant issues that directly impact the coupling of public health and environmental data as needed for environmental public health surveillance.

Restrictions on the use of PHI must be addressed to effectively utilize these data in environmental public health surveillance. The Health Insurance Portability and Accountability Act of 1996 (HIPAA) imposes numerous requirements on employer health plans concerning the use and disclosure of PHI, which includes virtually all individually identifiable health information held by a health plan — whether received in writing, in an electronic medium, or as an oral communication (http://www.hep-c-alert.org/links/hippa.html).

4.0 Evaluation

4.1 Technical Specifications and Feasibility for Using NASA Satellite Observations to Improve DSS

Within the respiratory health, birth defects and cancer teams, $PM_{2.5}$ has been identified as a vital surveillance data need. The requirement for linking with health data is that daily $PM_{2.5}$ estimates are available in a timely manner over the metropolitan Atlanta area. While use of the AQS $PM_{2.5}$ data alone could marginally meet HELIX-Atlanta specifications, there are only about five AQS sites in the Atlanta area, thus the spatial coverage is not ideal. Use of the MODIS satellite observations improves both the spatial resolution as well as the spatial coverage, providing a regional perspective that may prove helpful in identifying pollutant sources before they reach the Atlanta area. This is illustrated later in figs. 4-5. MODIS AOD observations are typically available within 3 hours after the overpass time; data download and processing to produce $PM_{2.5}$ estimates can be completed in minutes. Thus the data meet the timeliness requirement for the surveillance network.

4.2 History and limitations of derived $PM_{2.5}$ from MODIS AOD measurements

There have been a few recent studies investigating the potential for estimating $PM_{2.5}$ from MODIS satellite observations (Engel-Cox et al., 2004; Rush et al., 2004). Aerosol Optical Depth is only indirectly related to ground-based $PM_{2.5}$, and we rely on empirical correlations between the two quantities to enable us to estimate $PM_{2.5}$ from AOD. During the summer months, strong vertical mixing in the lower atmosphere (up to ~2 km above surface, called the boundary layer) results in a fairly even and consistent distribution of particulates, so the relationship between $PM_{2.5}$ and AOD is relatively strong. During the cool season, the relationship is weak. This has been attributed by other researchers to weaker boundary layer mixing or differences in $PM_{2.5}$ speciation between summer and winter (Engel-Cox et al., 2004; Rush et al., 2004). Results of Butler et al. (2003) for Atlanta show that sulfate is the dominant $PM_{2.5}$ constituent in summer, but in winter organic carbon is most prevalent. Nitrate, ammonium and metals proportions are higher in winter than in summer. It could also be that seasonal differences in the concentrations of larger particulates (> 2.5 microns) are responsible for the lower wintertime correlations.

The primary cause of the poor cool season correlations appears to be the weak boundary layer mixing. One factor supporting this hypothesis is that $PM_{2.5}$ speciation does not change drastically from summer to winter, yet the $PM_{2.5}$ – AOD correlations do.

Cloud cover prevents calculation of MODIS AOD. During summer in Atlanta, cloud cover is substantial during early afternoon when MODIS observations are made.

4.3 Utility of $PM_{2.5}$ data for HELIX-Atlanta

There are two general causes of asthma episodes -- the environment and genetic variants, each accounting for approximately 50% of the risk of the disease (Cookson and Moffatt, 2004). Children with asthma are particularly vulnerable to the adverse health effects of high levels of air pollution. Studies of children with asthma living in some of the most highly polluted regions of the world conclude that exposure to levels particulate matter (especially $PM_{2.5}$) regularly in excess of US EPA air quality standards significantly enhances the risk of respiratory symptoms, asthma medication use, and reduced lung function (Gent et al, 2003). These results are consistent with epidemiologic findings that show that metal composition of $PM_{2.5}$ influences the severity of allergic respiratory disease (Gavett et al, 2003). In fact, fine particles are most closely associated with such health effects as increased hospital admissions and emergency room visits for heart and lung disease, increased respiratory disease and symptoms such as asthma, decreased lung function, and even premature death. A study of half a million Kaiser Permanente members living in the San Joaquin Valley of California has reported that following wintertime episodes of high $PM_{2.5}$ and PM_{10} concentrations, and to a lesser extent carbon monoxide and nitrogen oxides, hospital admission rates and emergency room visits increased for patients who suffer from acute respiratory ailments such as asthma and bronchitis (Eeden, et al., 2002). Specifically, a 10 $\mu g/m^3$ rise in $PM_{2.5}$ concentrations increased the risk of having a more serious asthma attack the next day by 20% (Slaughter et al, 2003). The results of those studies were the rationale for choosing $PM_{2.5}$ as the environmental hazard for HELIX-Atlanta.

The EPA AQS network provides daily $PM_{2.5}$ measurements for different stations around the country with different reporting frequencies. However, those stations are concentrated in the metropolitan areas within the states with poor spatial coverage elsewhere. On the other hand, estimated $PM_{2.5}$ daily concentrations derived from MODIS observations can provide a very good spatial coverage with a reasonable resolution that compensate for the poor coverage of the AQS data, especially in the rural areas under cloud free conditions.

4.4 Design and Implementation of Derived $PM_{2.5}$

4.4.1 Data Availability and Retrieval

Daily Level 2 MODIS aerosol products for MODIS-Terra (AM overpass) and MODIS-Aqua (PM overpass) for the Atlanta region were downloaded from the NASA Data Gateway and processed. Terra data were obtained starting in June 2000 and Aqua data starting in July 2002, ending in December 2003 for both. The specific NASA data products are referred to as MOD04_L2 for Terra and MYD04_L2 for Aqua. The overpass times for Terra and Aqua are approximately 11:45 and 13:45 local standard time, respectively.

4.4.2 Development of Regression Models

AQS $PM_{2.5}$ data (in $\mu g/m^3$) were obtained from the U.S. EPA for the following sites in the 5-county Atlanta area:

South Dekalb	Daily	2000-2003
South Dekalb	Hourly	2002-2003
Doraville	Daily	2000-2003
East Rivers	Daily	2000-2003
Gwinnett Tech	Hourly	2003

For the daily AQS sites, $PM_{2.5}$ values represent a midnight-midnight local time mean. For the hourly South Dekalb and Gwinnett Tech sites, daily average $PM_{2.5}$ values were determined by averaging the 24 hourly readings from midnight to midnight local time. Regression equations have been developed using the 2002-2003 MODIS observations to evaluate the relationship between satellite-derived AOD and $PM_{2.5}$ measurements collected at the five AQS monitoring stations in the Atlanta study area. In order to use the MODIS AOD observations to estimate ground-level $PM_{2.5}$, regression models were established separately for the Terra and Aqua

MODIS observations. First, AOD observations corresponding to the locations of the AQS sites were extracted from the MODIS data files. This was done by selecting any AOD observations located within a 10 x 10 km box centered at the site location. If more than one MODIS observation fell within the box, the values were averaged to give the AOD value for the site. Linear correlation coefficients were then calculated on a monthly basis for each sensor, using all of the daily paired AOD - $PM_{2.5}$ observations for the month. Table 1 summarizes the monthly correlations. This analysis revealed that the AOD - PM2.5 relationship is generally weak during the cool season (October – March) and relatively strong during the warm season (April - September). This is very consistent with the results shown in Rush et al. (2004). Consequently, we grouped the data for April through September for each year and determined correlation coefficients and regression equations for each sensor. These are also shown in Table 1.

	2000	2001	2002		2003	
	Terra	Terra	Terra	Aqua	Terra	Aqua
January		0.062	0.121		0.036	0.432
February		0.553	0.475		0.728	0.198
March		0.038	-0.015		0.457	-0.133
April		0.511	0.326		0.806	0.397
May		0.431	0.269		0.328	0.469
June	-0.140		0.420		0.452	0.759
July	0.676	0.709	0.732	0.705	0.481	0.812
August	0.756	0.640	0.446	0.058	0.409	0.824
September	0.758	0.707	0.415	0.341	0.652	0.591
October	0.741	0.295	0.171	0.658	0.225	0.359
November	0.927	0.372	0.100	-0.077	0.052	0.164
December	0.181	0.234	0.224	-0.466	-0.411	0.151
April - September	0.579 (June - Sept.)	0.643	0.559	0.401 (July - Sept.)	0.661	0.727
Regression coefficients, April - September:						
Intercept	11.29	11.69	8.88	11.40	8.85	6.47
Slope	15.88	19.33	17.16	9.41	18.57	18.39

Table 1. Linear correlation coefficients by month and sensor, and regression coefficients for April-September for each year and sensor. $PM_{2.5}$ is the dependent variable and AOD is the dependent variable: $PM_{2.5}$ = Slope*AOD + Intercept

4.4.3 Data Transfer for HELIX-Atlanta

The two main sources of the environmental data for HELIX-Atlanta are the EPA AQS and the NASA EOS Data Gateway. The EPA AQS network provides daily $PM_{2.5}$ measurements and the NASA MODIS data provide Aerosol Optical Depth, which is used to derive $PM_{2.5}$ concentrations using the previously discussed regression equations. Those data sets can be downloaded off the EPA and NASA EOS Gateway websites. The environmental data (i.e. $PM_{2.5}$) will then be linked to the health data (i.e. acute asthma data), which is provided by Kaiser Permanente. KP will send the health data to NASA by overnight mail on CD media and the health data will be encrypted. The linked data that includes individual health information will be sent back from NASA to KP for further analysis on CD after applying the encryption procedure. Also, the linked data will be aggregated into surfaces of 10 km by 10 km resolution that will protect the privacy and confidentiality of KP health members. The aggregated data set will be sent back to the CDC and published for the public. Creating the daily spatial surfaces and tabular data output will take around 30 minutes to finish 365 surfaces of the year. Figure 2 shows a schematic of the data transfer process for HELIX-Atlanta.

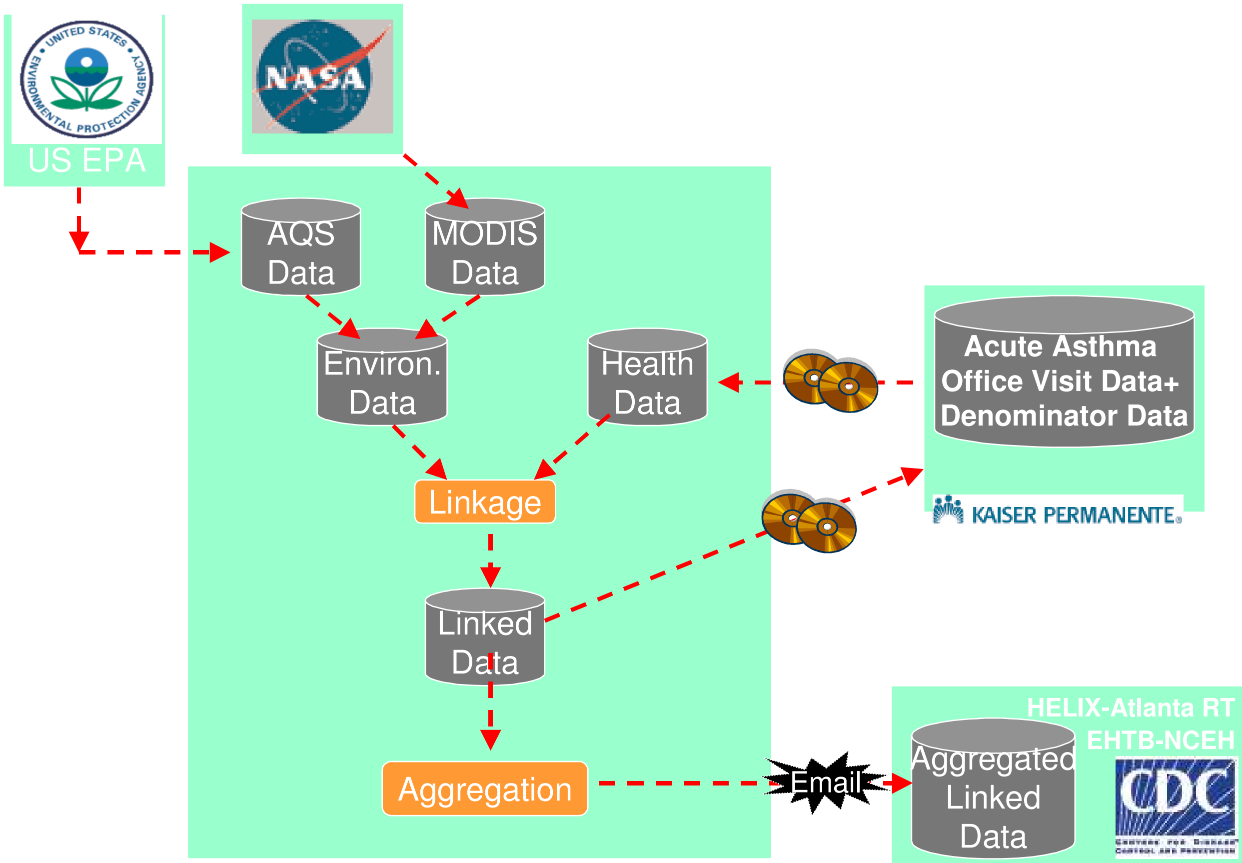


Figure 2. Schematic of the data transfer process for HELIX-Atlanta.

5.0 Verification and Validation

5.1 How do Derived $PM_{2.5}$ Values from Satellite Observations Meet Stated Specifications and Functional Requirements?

In the existing configuration, five AQS observation locations provided $PM_{2.5}$ data over the five county metropolitan Atlanta area. The requirement for HELIX-Atlanta is for daily, year-round estimates of $PM_{2.5}$ over the Atlanta area at a spatial resolution of ~10 km. This requirement can be met by creating surfaces of the daily values from the AQS observations from the five stations in the Atlanta area. During the cool season (October – March), the AQS data will be used alone to produce the spatial $PM_{2.5}$ product. We have verified that the MODIS AOD observations from the Terra and Aqua satellites can supplement the AQS data during the warm season (April-September). The NASA satellite observations will significantly enhance the spatial coverage and temporal frequency of the assimilated $PM_{2.5}$ product. In the following section we provide a comparison of the enhanced $PM_{2.5}$ product with an AQS-only product during the warm season.

6.0 Benchmarking

6.1 Benchmark $PM_{2.5}$ Values Derived from Satellite Observations Against AQS With and Without Quality Control

A quality control (QC) procedure for eliminating spurious measurements of $PM_{2.5}$ from a network of ground observations has been developed. The procedure utilizes observations from surrounding sites to determine whether a given measurement is acceptable or is considered erroneous and thus eliminated from further analysis. The filtered EPA AQS observations have then been used to generate spatially continuous ground-level $PM_{2.5}$ surfaces for the years 2000-2003 in order to estimate exposure to the local population using the B-Spline surfacing technique. The B-Spline technique handles data density and distribution issues better than most algorithms. It is also resistant to numerical artifacts and introduction of spurious spatial wavelengths and is relatively robust when the data contain spurious values. Figure 3 shows an example how using the QC procedure enhanced the output $PM_{2.5}$ surfaces by preventing any unrealistic ripples to be formed within the surface. The green arrow indicates the location of the anomalous value.

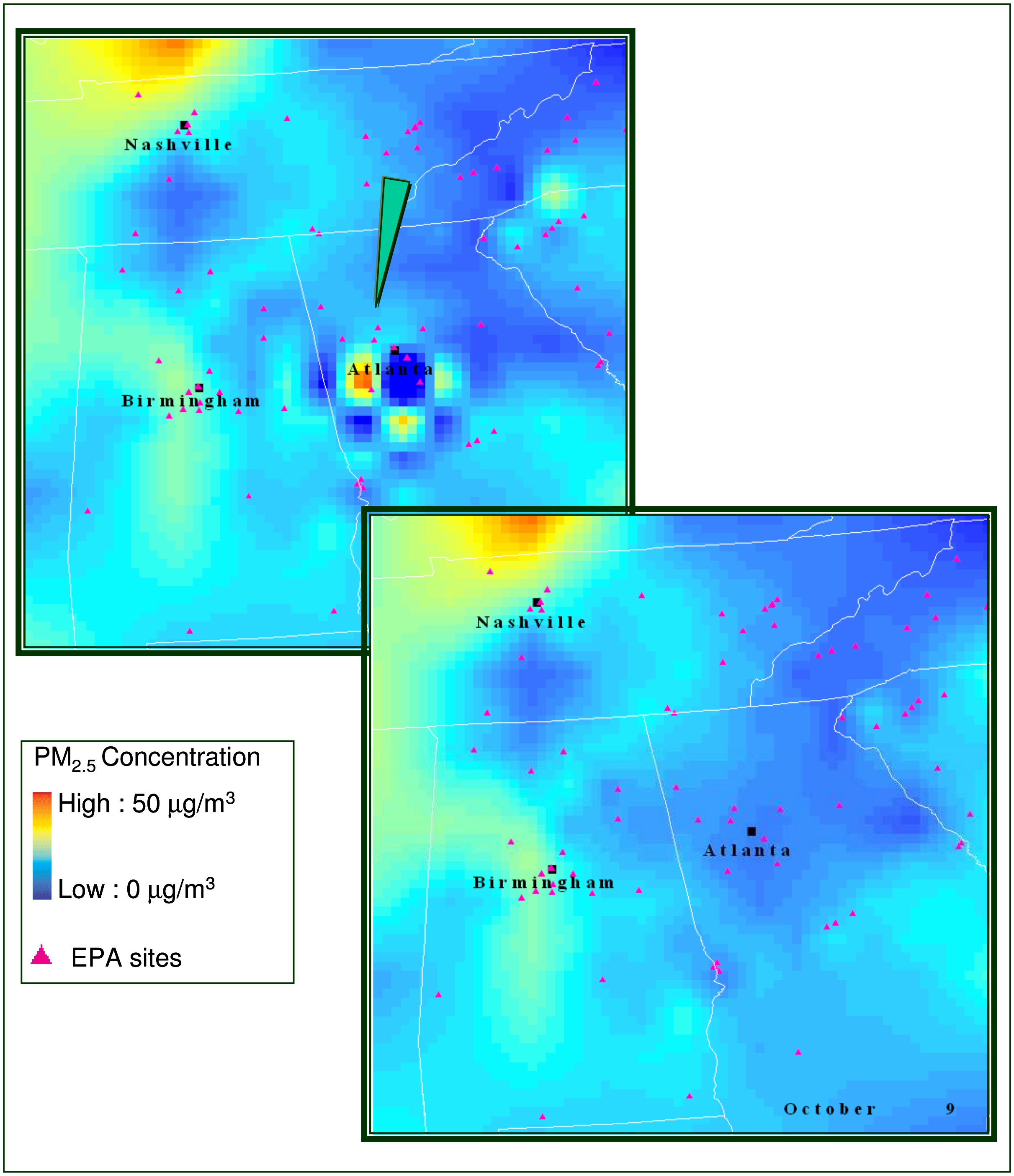
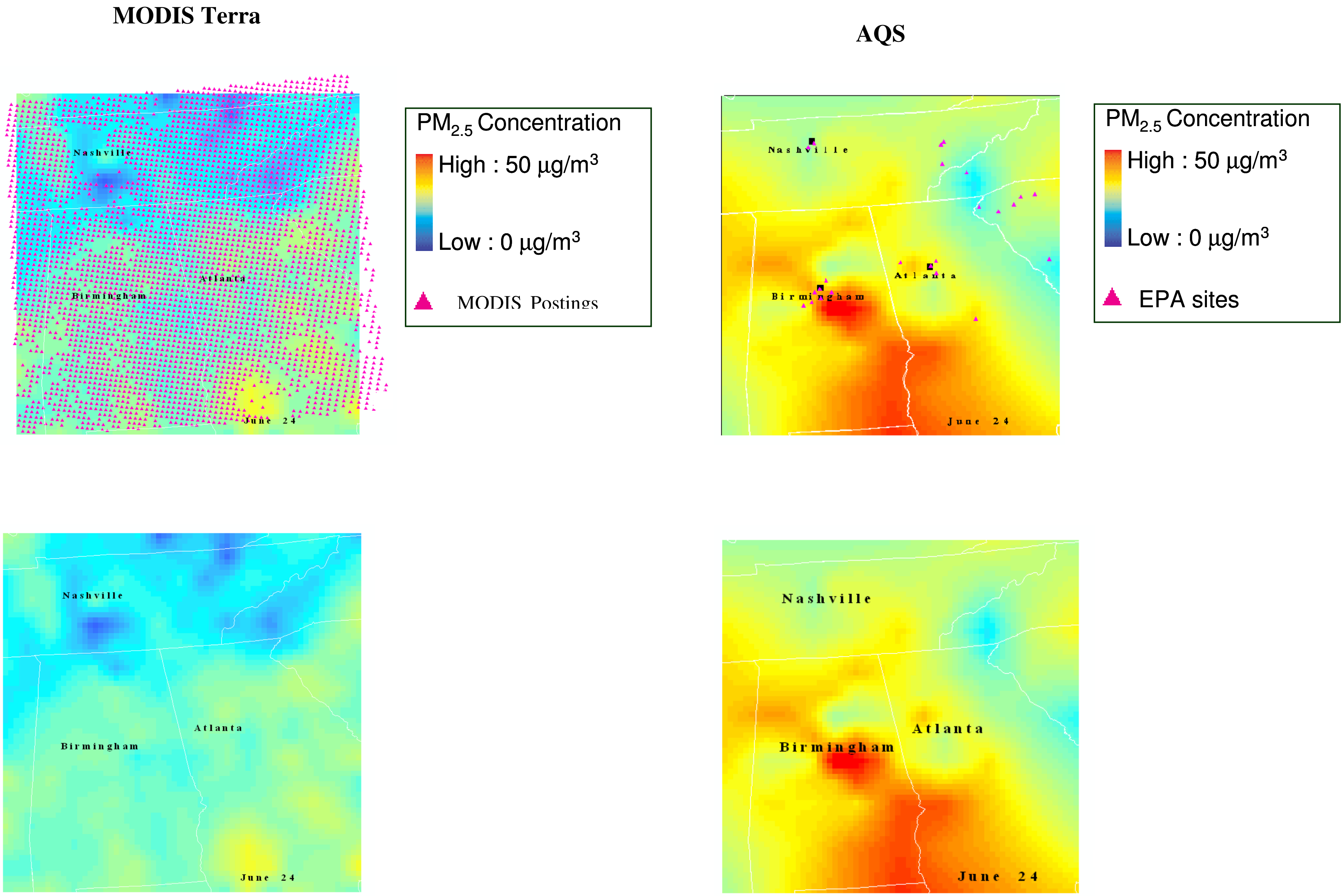


Figure 3. $PM_{2.5}$ B-Spline surfaces with and without quality control procedures using data from the EPA AQS network for October 9, 2003. The green arrow indicates the location of a value believed to be anomalous.

6.2 Enhanced HELIX-Atlanta DSS Performance

The EPA AQS network provides daily $PM_{2.5}$ measurements concentrated in the metropolitan areas and poor spatial coverage elsewhere. The coverage of NASA MODIS satellite observations can compensate for this problem, filling in the cloud-free areas with no $PM_{2.5}$ coverage from the AQS network. Poor spatial coverage will result in $PM_{2.5}$ estimations with high level of uncertainty in the areas devoid of observations.

A visual comparison between the B-Spline generated surfaces using the AQS data versus MODIS observations shows that the difference in the spatial coverage can result in surfaces that are quite different (Figures 4 and 5). The underlying concepts of the extrapolating surfacing techniques could be the main reason for that. The lack of observations in some areas within the surface (i.e. no boundary controls) as in the case of the AQS data can unreasonably cause the estimated values of the surface to go to the maximum or minimum possible value in the regions with no observations as shown in Figure 4b, which shows a generated B-Spline surface using EPA AQS data. On the other hand, Figure 4a shows that the good coverage of MODIS observations gave good $PM_{2.5}$ representation in the areas of the surface with no AQS observations preventing the red band shown in Figure 4b which showed an unrealistic increase in $PM_{2.5}$ concentrations away from the metropolitan areas (south part of the surface) in rural areas where there are no observations. Figure 5 shows another example of B-Spline surfaces using MODIS observations versus AQS data on August 18, 2003. On this day, we also can see that the poor spatial coverage of the AQS data caused an unreasonable increase in the $PM_{2.5}$ concentrations away from the metropolitan areas (southeast and northwest corners of the surface).



(a) (b)

Figure 4. B-Spline $PM_{2.5}$ surfaces for June 24, 2003 using (a) MODIS observations (b) AQS data

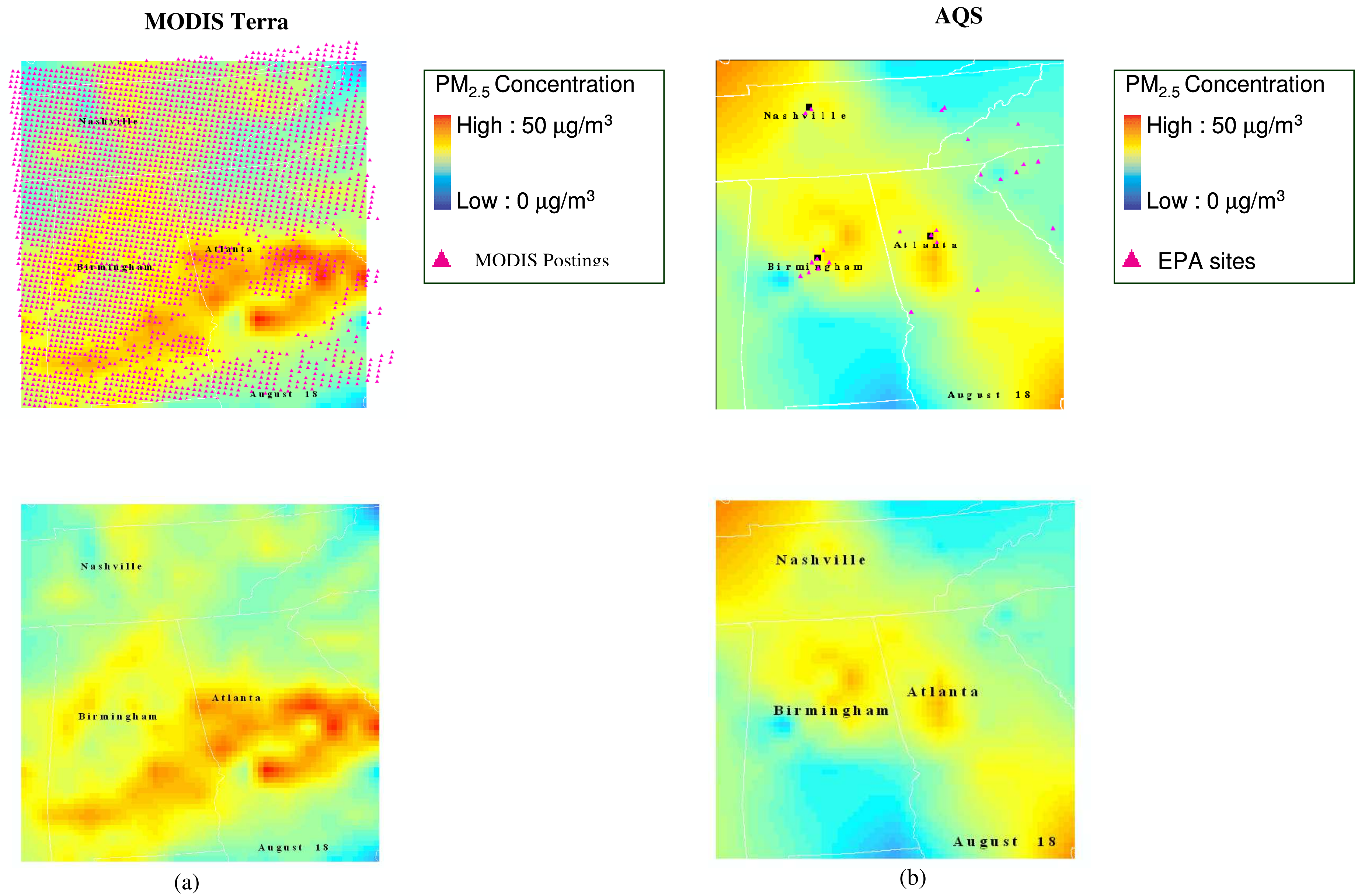


Figure 5. B-Spline $PM_{2.5}$ surfaces for August 18, 2003 using (a) MODIS observations (b) AQS data

7.0 Gaps in Benchmarking

- MODIS $PM_{2.5}$ estimates have not yet been integrated with the AQS data to produce the ultimate blended $PM_{2.5}$ maps.
- The spatial density of AQS data is not very high on most days, although it is better in the 5-county metropolitan Atlanta study area.
- MODIS satellite observations coverage is incomplete due to cloud cover.
- MODIS estimation of $PM_{2.5}$ during the cool season (October through March) is problematic with the approach used during the warm season. This result is consistent with previous research results and has been attributed by other researchers to weaker boundary layer mixing or differences in $PM_{2.5}$ speciation between summer and winter.
- A more robust evaluation of the enhancement of the HELIX-Atlanta DSS is needed when the DSS is fully operational.

7.1 Lessons Learned

- Quality control of $PM_{2.5}$ data is vital where spatial surfaces are to be created from the irregular point observations.
- Use of earth science satellite observations along with surface observations of $PM_{2.5}$ will provide a much more accurate regional representation of $PM_{2.5}$ than either data set alone would allow.

7.2 Recommendations

It is recommended that, in order to produce robust daily regional $PM_{2.5}$ maps, MODIS satellite observations be blended with surface observations. Care must be taken in this process to eliminate erroneous observations. Also, biases in the satellite observations with respect to ground observations need to be removed on a daily basis.

Other NASA satellite derived data products, including land surface temperature data, should be evaluated in HELIX-Atlanta to determine its potential benefit for EPHT.

8.0 Appendices

8.1 Acronyms

AOD – Aerosol Optical Depth

AQS – Air Quality System

BAA – Business Associate Agreement

CDC – Centers for Disease Control and Prevention

DSS – Decision Support System

DST – Decision Support Tool

EPA – Environmental Protection Agency

EPD – Environmental Protection Division

EPHT – Environmental Public Health Tracking

EPHTN – Environmental Public Health Tracking Network

GEOSS – Global Earth Observation System of Systems

GIS – Geographic Information System

HELIX-Atlanta - Health and Environment Linked for Information Exchange in Atlanta

HIPAA – Health Insurance Portability and Accountability Act

ISS – Integrated System Solution

IWGEO – Interagency Working Group on Earth Observations

KP-GA – Kaiser Foundation Health Plan of Georgia, Inc.

MSFC – Marshall Space Flight Center

MODIS - Moderate Resolution Imaging Spectroradiometer

NAAQS – National Ambient Air Quality Standards

NASA – National Aeronautics and Space Administration

NCEH – National Center for Environmental Health

NEPA - National Environmental Policy Act

PHI – Protected Health Information

PHIN – Public Health Information Network

$PM_{2.5}$ – Particulate Matter with diameter of 2.5 microns or less

V&V – Verification and Validation

WRI - World Research Institute

8.2 References

Cookson, W. and M. Moffatt, Making Sense of Asthma Genes. *New Engl. J. Med*. 2004;351:1794-1796.

Eeden, V. D. et al., Particulate Air Pollution and Morbidity in the California Central Valley: A High Particulate Pollution Region. Final Report to the California Air Resources Board, Contract 97-303, July 12, 2002.

Gent, J. F. et al., Association of Low-Level Ozone and Fine Particles with Respiratory Symptoms in Children with Asthma. *J. Am. Med. Assoc*. 2003; 290: 1859-1867.

Gavett, S. H. et al., Metal Composition of Ambient $PM_{2.5}$ Influences Severity of Allergic Airways Disease in Mice. *Envir. H. Persp*. 2003; 111(12) 1471-1477.

GEOSS. Global Earth Observation System of Systems GEOSS, 10-Year Implementation Plan Reference Document, Group on Earth Observations. ESA Publications Division, Noordwijk, The Netherlands, 210 pp, 2005.

NASA Langley Research Center, Stennis Research Center, Goddard Space Flight Center, The Application of Satellite Data for Forecasting Particle Pollution, 25 pp, 2003.

McGeehin, M. A., J. R. Qualters, and A. S. Niskar; 2004; National Environmental Public Health Tracking Program: Bridging the Information Gap; Environmental Health Perspectives, vol. 112, no. 14, October 2004, pp 1409 – 1413.

NASA. Earth-Sun System Applied Sciences Program, Public Health Program Element, FY2005-2009 Plan. NASA Science Mission Directorate, Version 1.1, March 16, 2005. NASA Headquarters, Washington, DC, 27 pp, 2005.

Slaughter, J. C. et al., Effects of ambient air pollution on symptom severity and medication use in children with asthma. Ann. Allergy Asthma Immunol. 2003; 91(4): 346-353.

World Resources Institute (WRI), 1999. Health Effects of Air Pollution. http://www.wri.org/wr-98-99/airpoll.htm, accessed November 5, 2003.